WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 13, 2023 – 12:50 am GMT

|		
PDB ID	:	4BQK
Title	:	rice importin_alpha : VirD2NLS complex
Authors	:	Chang, C.-W.; Counago, R.M.; Williams, S.J.; Kobe, B.
Deposited on	:	2013-05-31
Resolution	:	2.00 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
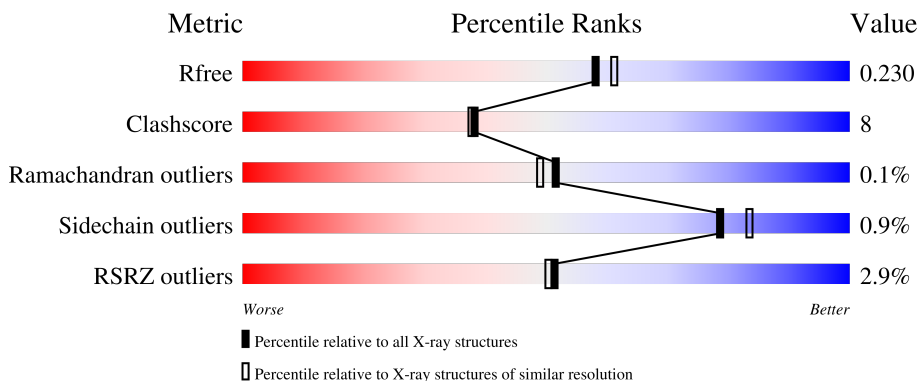
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	456	2% · 82% · 10% · 7%
1	B	456	3% · 81% · 11% · 7%
2	C	20	10% · 55% · 5% · 40%
2	D	20	55% · 15% · 30%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PEG	A	1495	-	-	X	-
3	PEG	A	1496	-	-	X	-
3	PEG	A	1498	-	-	X	-
3	PEG	B	1495	-	-	-	X
3	PEG	B	1496	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 13793 atoms, of which 6630 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called IMPORTIN SUBUNIT ALPHA-1A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	424	Total 6397	C 2048	H 3156	N 549	O 629	S 15	0	0	0
1	B	422	Total 6370	C 2043	H 3139	N 547	O 626	S 15	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	71	GLY	-	expression tag	UNP Q71VM4
A	72	SER	-	expression tag	UNP Q71VM4
B	71	GLY	-	expression tag	UNP Q71VM4
B	72	SER	-	expression tag	UNP Q71VM4

- Molecule 2 is a protein called T-DNA BORDER ENDONUCLEASE VIRD2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total 228	C 63	H 118	N 29	O 18	0	0	0
2	D	14	Total 250	C 70	H 127	N 31	O 22	0	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 17	C 4	H 10	O 3	0	0
3	A	1	Total 17	C 4	H 10	O 3	0	0
3	A	1	Total 17	C 4	H 10	O 3	0	0
3	A	1	Total 17	C 4	H 10	O 3	0	0
3	B	1	Total 17	C 4	H 10	O 3	0	0
3	B	1	Total 17	C 4	H 10	O 3	0	0
3	B	1	Total 17	C 4	H 10	O 3	0	0
3	B	1	Total 17	C 4	H 10	O 3	0	0
3	B	1	Total 17	C 4	H 10	O 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	229	Total 229	O 229	0	0
4	B	157	Total 157	O 157	0	0
4	C	5	Total 5	O 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	4	Total O 4 4		0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2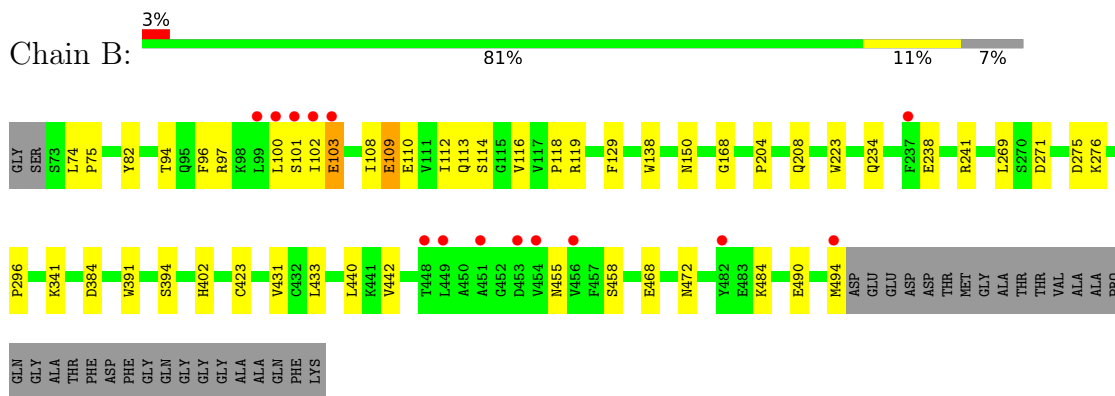 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: IMPORTIN SUBUNIT ALPHA-1A
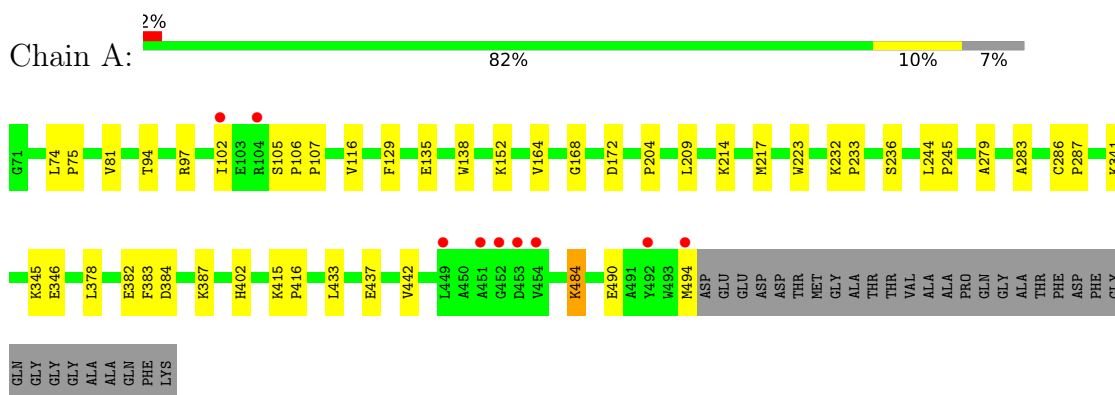
• Molecule 1: IMPORTIN SUBUNIT ALPHA-1A

• Molecule 2: T-DNA BORDER ENDONUCLEASE VIRD2

• Molecule 2: T-DNA BORDER ENDONUCLEASE VIRD2

L415 R420 E421 ASP ASP ASP GLY GLU PRO S428 R432 E433 R434

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.80Å 140.95Å 72.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.72 – 2.00 19.72 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.6 (19.72-2.00) 99.4 (19.72-2.00)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.01 (at 1.99Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8.1_1168)	Depositor
R, R_{free}	0.196 , 0.224 0.227 , 0.230	Depositor DCC
R_{free} test set	4244 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	19.4	Xtriage
Anisotropy	0.324	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 29.0	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.159 for h,-k,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	13793	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.51% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PEG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.63	0/3296	0.66	0/4488
1	B	0.60	0/3286	0.64	0/4475
2	C	0.57	0/109	0.67	0/138
2	D	0.69	0/122	0.89	0/155
All	All	0.62	0/6813	0.65	0/9256

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3241	3156	3274	49	18
1	B	3231	3139	3266	50	18
2	C	110	118	118	1	0
2	D	123	127	125	2	0
3	A	28	40	40	21	0
3	B	35	50	50	10	0
4	A	229	0	0	4	0
4	B	157	0	0	4	0
4	C	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	4	0	0	0	0
All	All	7163	6630	6873	103	18

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

The worst 5 of 103 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:101:SER:N	1:B:102:ILE:HA	1.96	0.81
1:B:103:GLU:HB3	1:B:108:ILE:HG21	1.65	0.78
3:B:1495:PEG:H21	4:B:2156:HOH:O	1.82	0.77
1:A:217:MET:HB2	3:A:1496:PEG:H32	1.67	0.76
3:B:1495:PEG:C2	4:B:2156:HOH:O	2.34	0.75

The worst 5 of 18 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:75:PRO:HD3	1:B:114:SER:O[2_655]	0.38	1.22
1:A:75:PRO:HB3	1:B:82:TYR:OH[2_655]	0.41	1.19
1:A:74:LEU:HB3	1:B:114:SER:OG[2_655]	0.42	1.18
1:A:75:PRO:HG3	1:B:118:PRO:CD[2_655]	0.45	1.15
1:A:74:LEU:CD2	1:B:110:GLU:HB3[2_655]	0.56	1.04

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	422/456 (92%)	418 (99%)	4 (1%)	0	100	100
1	B	420/456 (92%)	414 (99%)	5 (1%)	1 (0%)	47	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	8/20 (40%)	8 (100%)	0	0	100	100
2	D	10/20 (50%)	9 (90%)	1 (10%)	0	100	100
All	All	860/952 (90%)	849 (99%)	10 (1%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	103	GLU

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	361/381 (95%)	358 (99%)	3 (1%)	81	86
1	B	360/381 (94%)	356 (99%)	4 (1%)	73	78
2	C	12/19 (63%)	12 (100%)	0	100	100
2	D	13/19 (68%)	13 (100%)	0	100	100
All	All	746/800 (93%)	739 (99%)	7 (1%)	78	83

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	109	GLU
1	B	129	PHE
1	B	423	CYS
1	B	208	GLN
1	A	484	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	150	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	220	ASN
1	B	254	ASN
1	B	252	HIS
1	A	252	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
3	PEG	A	1498	-	6,6,6	0.46	0	5,5,5	0.54	0
3	PEG	B	1496	-	6,6,6	0.52	0	5,5,5	0.82	0
3	PEG	A	1495	-	6,6,6	0.49	0	5,5,5	1.28	1 (20%)
3	PEG	A	1496	-	6,6,6	0.49	0	5,5,5	0.70	0
3	PEG	B	1499	-	6,6,6	0.48	0	5,5,5	0.58	0
3	PEG	B	1497	-	6,6,6	0.62	0	5,5,5	0.58	0
3	PEG	B	1498	-	6,6,6	0.51	0	5,5,5	0.60	0
3	PEG	A	1497	-	6,6,6	0.59	0	5,5,5	0.82	0
3	PEG	B	1495	-	6,6,6	0.50	0	5,5,5	0.83	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	PEG	A	1498	-	-	3/4/4/4	-
3	PEG	B	1496	-	-	1/4/4/4	-
3	PEG	A	1495	-	-	2/4/4/4	-
3	PEG	A	1496	-	-	2/4/4/4	-
3	PEG	B	1499	-	-	0/4/4/4	-
3	PEG	B	1497	-	-	2/4/4/4	-
3	PEG	B	1498	-	-	4/4/4/4	-
3	PEG	A	1497	-	-	4/4/4/4	-
3	PEG	B	1495	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1495	PEG	C3-O2-C2	2.63	124.68	113.29

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1497	PEG	O2-C3-C4-O4
3	A	1498	PEG	O1-C1-C2-O2
3	B	1496	PEG	O1-C1-C2-O2
3	B	1497	PEG	C1-C2-O2-C3
3	A	1495	PEG	O2-C3-C4-O4

There are no ring outliers.

6 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1498	PEG	4	0
3	B	1496	PEG	5	0
3	A	1495	PEG	6	0
3	A	1496	PEG	11	0
3	B	1498	PEG	2	0
3	B	1495	PEG	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	424/456 (92%)	0.45	9 (2%)	63	62	13, 23, 47, 98	0
1	B	422/456 (92%)	0.55	14 (3%)	46	45	16, 25, 58, 97	0
2	C	12/20 (60%)	0.67	2 (16%)	1	1	24, 39, 47, 52	0
2	D	14/20 (70%)	0.57	0	100	100	23, 38, 52, 75	0
All	All	872/952 (91%)	0.50	25 (2%)	51	50	13, 24, 54, 98	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	102	ILE	7.3
1	A	494	MET	6.8
1	B	103	GLU	6.6
1	B	99	LEU	6.1
1	B	451	ALA	5.8

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PEG	B	1495	7/7	0.64	0.41	26,64,83,83	0
3	PEG	A	1496	7/7	0.72	0.38	32,48,58,58	0
3	PEG	A	1498	7/7	0.73	0.35	28,45,56,56	0
3	PEG	A	1495	7/7	0.74	0.33	25,38,53,55	0
3	PEG	A	1497	7/7	0.76	0.32	54,65,73,80	0
3	PEG	B	1498	7/7	0.79	0.30	30,51,67,67	0
3	PEG	B	1497	7/7	0.80	0.23	43,52,61,62	0
3	PEG	B	1496	7/7	0.86	0.41	27,48,68,68	0
3	PEG	B	1499	7/7	0.86	0.20	29,42,51,53	0

6.5 Other polymers ⓘ

There are no such residues in this entry.